Gene 642 (2018) 64–73

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Transcriptome – Scale characterization of salt responsive bean TCP transcription factors

Emre İlhan[a,*], İlker Büyük[b,c], Behcet İnal[d]

[a] Depart. of Molecular Bio. and Genetics, Erzurum Technical University, Erzurum, Turkey
[b] Depart. of Biology, Ankara University, Ankara, Turkey
[c] Depart. of Plant, Soil and Microbial Sciences, Michigan State University, East Lansing, MI 48824, USA
[d] Depart. of Agricultural Biotechnology, Siirt University, Siirt, Turkey

ARTICLE INFO

Keywords:
Phaseolus vulgaris
Gene expression
Salt stress
Transcription factor

ABSTRACT

TEOSINTE-BRANCHED1/CYCLOIDEA/PCF (TCP) proteins are important regulators of growth and developmental processes including branching, floral organ morphogenesis and leaf growth as well as stress response. This study identified 27 *TCP* genes of *Phaseolus vulgaris* (common bean), which were divided into three clusters based on phylogenetic relationship. In addition, this study showed that some of *TCP* genes such as *Pvul-TCP-4* and *Pvul-TCP-15* located on chromosomes 3 and 7, *Pvul-TCP-7* and *Pvul-TCP-20* located on chromosome 7 and 9, were segmentally duplicated. On the other hand, a total of 20 *Pvul-TCP* genes have predicted to be targeted by microRNAs (miRNA). Most of the miRNA-target genes were *Pvul-TCP-1*, *-11*, *-13* and *-27*, which were targeted by 13, 17, 22 and 13 plant miRNAs, respectively. miR319 was one of the highly represented regulatory miRNAs to target TCP transcripts. Promoter region analysis of TCP genes resulted that the GT-1 motif, which was related to salt stress, was found in 14 different *Pvul-TCP* genes. Expression profiling of 10 *Pvul-TCP* genes based on RNA-sequencing data further confirmed with quantitative real-time RT-PCR measurements identified that *Pvul-TCP* genes under salt stress are expressed in a cultivar- and tissue-specific manner.

1. Introduction

TEOSINTE-BRANCHED1/CYCLOIDEA/PCF (TCP) proteins are important regulators of growth and developmental processes such as shoot branching, floral organ morphogenesis, leaf growth, female and male gametophyte development, and circadian clock (Zhou et al., 2016). The TCP domain is vital for regulation of gene expression and protein – protein interactions (Manassero et al., 2013). On the other hand, some of the TCP domains contain a basic helix loop helix (bHLH). This bHLH differ from structure of bHLH transcription factors which is a DNA binding domain both in plants and animals (Cubas et al., 1999). TCPs consist of TEOSINTE BRANCHED (TB from maize) 1, CYCLOIDEA domains (CYC from *Antirrhinum majus*) and PROLIFERATING CELL FACTORS (PCFs from rice) (Doebley et al., 1997; Kosugi and Ohashi, 1997; Cubas et al., 1999; Luo et al., 1999). Additionally, TCP gene family was identified in the plants that have draft genome data. Until now, TCPs were found in *Arabidopsis* (23), rice (22), apple (52) (Xu et al., 2014), Gossypium raimondii (38) (Ma et al., 2014), sorghum (20) (Francis et al., 2016), strawberry (19) (Wei et al., 2016) and tomato (30) (Parapunova et al., 2014). TB1 and CYC serve as negative regulators for lateral branching and floral development respectively. PCFs encourage cell proliferation and organ growth (Takeda et al., 2003; Li, 2015). According to bioinformatics analysis, TCP proteins are divided into two classes; class I and class II. While class I TCP proteins have one subclass (TCP – P), class II TCPs are divided into two subclasses: ubiquitous CINCINNATA; CIN and angiosperm – specific CYC/TB1. The clade CIN genes have function in lateral organ development and the clade CYC/TB1 genes are basically important in lateral shoot or flower development. The members of class II TCPs have a conserved arginine – rich R domain with unknown function (Crawford et al., 2004; Ma et al., 2014). Another feature of *TCP* genes have targeted by miR319, which involves in regulation of leaf morphogenesis. In *Arabidopsis*, the genes such as *TCP2*, *TCP3*, *TCP4*, *TCP10* and *TCP24* are targeted by miR319 (Palatnik et al., 2003; Xu et al., 2014). Recent studies showed that the TCP gene family has gene duplications (Ma et al., 2014; Parapunova et al., 2014; Xu et al., 2014).

Salinity stress causes accumulation of toxic ions in cells. Since extreme salt affects nutrient and water uptake of plants, crop yields in agricultural production and plant growth are restricted. The plants respond to salt stress to minimize salt harm and re-built ion homeostasis

Abbreviations: TCP, TEOSINTE-BRANCHED1/CYCLOIDEA/PCF; miRNA, microRNAs; bHLH, basic helix loop helix; PCFs, PROLIFERATING CELL FACTORS

* Corresponding author.
E-mail address: emre.ilhan@erzurum.edu.tr (E. İlhan).

https://doi.org/10.1016/j.gene.2017.11.021
Received 15 March 2017; Received in revised form 22 October 2017; Accepted 7 November 2017
Available online 09 November 2017
0378-1119/

and adapt to salt stress by activating a complex network of genes upon exposure to stress. Thus, the genetically occurring variations such as wild, varieties, cultivars and landraces provide important advantages to understand response mechanisms of salt stress trough both molecular techniques and conventional breeding (Kamei et al., 2005; Hussein et al., 2014; Hossain et al., 2016).

In this study, bioinformatics approaches were used for the identification of new proteins and genes. Although omics technologies such as genomics, transcriptomics and metabolomics have become widespread, there is no study that aimed at genome-wide identification and expression analysis of *TCP* gene family in common bean. The expression analysis of randomly selected *TCP* genes were carried out using tolerant Yakutiye and sensitive Zulbiye genotypes of *Phaseolus vulgaris* under salt stress. The results of this study provide important findings for further breeding studies of *P. vulgaris* focusing on response to salt stress.

2. Materials and methods

2.1. Sequence retrieval of TCP genes

Sequences of TCP proteins of *Phaseolus vulgaris* were obtained from Phytozome database v11 (www.phytozome.net) using key searching with Pfam Accession Number (PF03634) downloaded from Pfam Database (http://pfam.xfam.org/). Also, protein sequences of *Zea mays*, *Antirrhinum majus* (AmCYC, AmCIN) and *Oryza sativa* (OsPCF) were obtained (Zhou et al., 2016). Twenty-seven genes identified in *P. vulgaris* genome were used in this study. Hidden Markov model (HMM, http://www.ebi.ac.uk) searches were performed against the *P. vulgaris* genome with default parameters (Goodstein et al., 2012). TCP proteins were also subjected to query in blastp (NCBI) for characterization of hypothetical proteins. Redundant sequences were removed by using the decrease redundancy tool (http://web.expasy.org/decrease_redundancy/). Non – redundant sequences were used to check the presence of TCP domains by using HMMER (http://www.ebi.ac.uk) and Pfam databases. Molecular weights, instability index and theoretical isoelectric points (pI) of identified proteins were calculated by using ProtParam Tool (http://web.expasy.org/protparam).

2.2. Structures and physical locations of Pvul-TCP genes, and detection of gene duplication events

Gene Structure Display Server v2.0 (GSDS, http://gsds.cbi.pku.edu.cn/) was used to determine the exon – intron sites of *Pvul-TCP* genes (Guo et al., 2007). Both genome sequences and the coding sequences were utilized for prediction of the positional information of the *Pvul-TCP* genes. Chromosomal locations, sizes (bp) and intron numbers were identified by using Phytozome database v11, The *Pvul-TCP* genes were plotted on all *P. vulgaris* chromosomes and pictured with MapChart (Voorrips, 2002). Gene duplication events among all putative *Pvul-TCP* genes were defined according to the following parameters: the alignment of the coding nucleotide sequences covered 70% of the longest genes and the amino acid identity between the sequences was > 70% (Gu et al., 2002; Yang et al., 2008).

To identify additional conserved motifs of Pvul-TCP proteins, we used the Multiple EM for Motif Elicitation tool (MEME v4.11.1; http://meme-suite.org/) (Bailey et al., 2006). The limits for minimum/maximum width and maximum number of motifs were set as 2, 50, and 20, respectively. Motif sites were between 2 and 300. Site distribution was adjusted as any number of repetitions. The identified conserved motifs were searched in the InterProscan database with default settings (Quevillon et al., 2005).

2.3. Phylogenetic analysis and sequence alignment

Phylogenetic analysis was performed using Neighbor-joining (NJ) algorithm with bootstrap value of 1000 replicates. ClustalW was used for the purpose of aligning the Pvul-TCP, 1 ZmTB1, 1 AmCYC, 1 AmCIN and 2 OsPCFs protein sequences with the gap opening and the gap extension penalties of 10 and 0.1, respectively (Thompson et al., 1997). MEGA6.0 (Tamura et al., 2011) was used to construct the phylogenetic tree, which was visualized by Interactive tree of life (iTOL; http://itol.embl.de/index.shtml) (Letunic and Bork, 2011).

2.4. Subcellular localization and promoter analysis of the Pvul-TCP family

The 5′ upstream region, including a 2 kb DNA sequence of each gene of the Pvul-TCP family, was analyzed using the plantCARE database (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) for a cis-element scan. The details of the 2 kb DNA sequence of each Pvul-TCP genes were given as Supplementary File S1. Subcellular localization of the Pvul-TCP family was predicted using the WoLF PSOPT (http://www.genscript.com/psort/wolf_psort.html) (Horton et al., 2007).

2.5. Ortholog relationships between common bean and other species

To obtain orthologous relationship among the chromosomes of common bean, soybean, *Medicago truncatula*, *Arabidopsis thaliana* and *Populus trichocarpa* species, TCP proteins of common bean were searched against amino acid sequences of *A. thaliana*, *M. truncatula*, soybean and *P. trichocarpa* (www.phytozome.net) using NCBI - BLASTP. Hits with $\leq 10–5$ and at least 60% coverage and 60% identity were identified. The comparative orthologous relationships of TCP genes among common bean, soybean, *A. thaliana*, *M. truncatula* and *P. trichocarpa* chromosomes were drawn using MapChart.

2.6. Annotation

The functional annotation of Pvul-TCP protein sequences were executed by using Blast2GO (http://www.blast2go.com) (Conesa et al., 2005). Molecular functions, cellular components and biological processes were obtained from output of the Blast2GO program.

2.7. miRNA prediction

The all known plant miRNA sequences were obtained from miRBase v21.0 (http://www.mirbase.org). All plant miRNAs and common bean *TCP* gene targets were aligned using the web-based psRNA Target Server (http://plantgrn.noble.org/psRNATarget) with default parameters. Alignments were calculated according to Zhang (2005). Computationally predicted miRNA targets were searched with BLASTX with $\leq 1e-10$ against common bean EST sequences at NCBI database for the identification and validation of putative gene homology.

2.8. Expression analysis of Pvul-TCP genes

RNA-sequencing reads with the following accession numbers; SRR957667 (control leaf), SRR958472 (salt-treated root), SRR958469 (control root), and SRR957668 (salt-treated leaf), which were reported by Hiz et al. (2014), were downloaded from Sequence Read Archive (SRA) and, were used to assess the expression patterns of common bean Pvul-TCP genes. After raw reads were downloaded as a single SRA file, it was divided into two paired-end files and converted to “fastq” format by using the NCBI SRA Toolkit's fastq-dump command. The quality of fastq files were evaluated with FASTQC and CLC Genomics Workbench 7.0 and the low-quality reads were eliminated. After the sequence quality control, the reads were aligned to *Phaseolus vulgaris* genome (v1.0) downloaded from PHYTOZOME v11 database. Uniquely mapped reads were used in order to determine the expression level. To normalize the gene expression values, the reads per kilobase of exon model per million mapped reads (RPKM) algorithm were used in this study (Mortazavi et al., 2008). Finally, the heat maps of hierarchical